Assessing Honesty-Humility: Theoretical and Practical Considerations of Traditional and Game-Based Assessments

Barends, A.J.

2020

document version
Publisher's PDF, also known as Version of record

Chapter 3. Gamified Personality Assessment: Virtual Behavior Cues of Honesty-Humility

This chapter is based on Barends, A. J., & De Vries, R. E., & Van Vugt, M. (2019). Gamified personality assessment: Virtual behavior cues of honesty-humility. *Zeitschrift für Psychologie, 227*(3)*,* 207-217.

3. Gamified Personality Assessment: Virtual Behavior Cues of Honesty-Humility

Abstract

Unobtrusive behavioral cues of personality traits can be found in physical and virtual environments (e.g., office environments and social media profiles), but detecting and coding such cues are a painstaking effort, and therefore impractical for research purposes. Yet, measuring people's choices in a virtual, gamified environment may offer a suitable substitute. Currently, it is unknown whether Honesty-Humility can also be assessed in a virtual environment. In two studies, we demonstrate that Honesty-Humility can be inferred with at least modest validity from virtual behavior cues. In a third study, we tested the fakeability of the virtual cues. This study found that even under faking instructions the virtual cues were related to Honesty-Humility, however, the virtual cues were just as fakeable as self-reported Honesty-Humility. Our results imply that virtual cues can be incorporated in assessment games to measure personality. Future research may investigate whether the identified virtual cues are able to predict important Honesty-Humility related outcomes.

Keywords: Personality, Honesty-Humility, gamification, virtual cues, faking.

3. Gamified Personality Assessment: Virtual Behavior Cues of Honesty-Humility

Features in the physical and virtual environments can tell a lot about their occupants. That is, personality seems to influence how people choose to dress themselves, how they customize their home environment, and what kind of social media profiles they have. The observable results of these behaviors serve as potential cues of an individual's personality. More precisely, behavioral personality cues (hereafter personality cues for short) are the observable features in physical and virtual environments resulting from behavior that reflect a person's personality. Earlier work has found that valid personality cues are available in many environmental features such as physical appearance (Naumann, Vazire, Rentfrow, & Gosling, 2009), Twitter profiles (Quercia, Kosinski, Stillwell, & Crowcroft, 2011), living spaces (Gosling, Ko, Mannarelli, & Morris, 2002), and office spaces (Gosling et al., 2002).

Although personality cues have been investigated in the context of the five-factor approaches (e.g., Gosling et al., 2002; Naumann et al., 2009), no such cues have been investigated yet for the HEXACO's Honesty-Humility trait. Honesty-Humility is not represented well in the five-factor model (Ashton, Lee, & De Vries, 2014). Consequently, it is unknown whether Honesty-Humility can be adequately inferred from personality cues. Honesty-Humility may be more difficult to infer from personality cues because people may be wary about leaving behind cues that may suggest they are low on this trait. However, indirect evidence suggests that there are valid cues for this trait. For instance, studies have shown that Narcissism, a trait closely related to low Honesty-Humility (Lee & Ashton, 2005), can be inferred from physical appearance (e.g., expensive clothing; Vazire, Naumann, Rentfrow, & Gosling, 2008) and signatures (Mailhos, Buunk, & Cabana, 2016). Together, these two studies imply that it is likely that Honesty-Humility can also be validly inferred from personality cues.

Research into personality cues is very labor intensive because of the vast number of cues available in the environment. Therefore, an alternative is to investigate virtual variants of the personality cues (i.e., virtual cues) in a gamified setting. The advantage of this approach is that researchers have much more control over the range of cues (i.e., a variety of topics not normally present in a person's environment can be covered), the number of cues, and the scoring of cues (i.e., in a virtual environment these can be scored automatically). Thus, the objectives of this study are to explore a range of potential virtual cues of Honesty-Humility, to develop and test the reliability and validity of a gamified assessment instrument based on these virtual cues, and to test its practical utility for personnel selection.

Personality Cues

Two general types of personality cues can be distinguished: identity claims and behavioral residues (Gosling et al., 2002; Graham, Sandy, & Gosling, 2011).[1] Identity claims are personality cues that reflect an individual's deliberate attempts to communicate one's personal taste or identity (e.g., dyeing one's hair blue to be perceived as rebellious), whereas behavioral residues are the traces left behind in the environment (e.g., physical appearance, physical environment, and social media profiles) from behavior that reflects personality (e.g., a messy office is likely the result of low Conscientiousness). Both types of cues contain valid personality information (e.g., Gosling et al., 2002; Naumann et al., 2009).

An obstacle for investigating personality cues is that it is very labor intensive to code the vast number of personality cues in environments. This is best illustrated in the Personal Living Space Cue Inventory (PLSCI; Gosling, Craik, Martin, & Pryor, 2005). Gosling et al. used the PLSCI to code rooms on 42 global cue categories, a task which took single raters about 45 minutes for one room. Additionally, about 100 cues could not be coded, which led Gosling et al. to note that no coding scheme would suffice in capturing all available cues.

Virtual Cues

Gamification may make it possible to use personality cues for assessment purposes in the form of virtual cues. Gamification is defined as the use of game design elements in non-game contexts (Deterding, Dixon, Khaled, & Nacke, 2011). Various studies have demonstrated that someone's behavior in a game reflects some aspect of their personality (e.g., Triberti, Villani, & Riva, 2015; Worth & Book, 2014; 2015). Interpreting these findings in terms of the Situation-Trait-Outcome Activation model (De Vries, Tybur, Pollet, & Van Vugt, 2016) leads us to expect why this is the case; that is, people select situations that reflect their personality, in these situations traits are activated, and each time a trait is activated, positive and/or negative outcomes result. With respect to situation activation (e.g., selection of situations to fit one's personality), Worth and Book (2014) found in the game World of Warcraft that people low in Honesty-Humility were more likely to engage in player-versus-player combat, whereas people high in Conscientiousness were more likely to spend time collecting resources and crafting in-game items (e.g., weapons). With respect to trait activation (i.e., the expression of a trait in a trait-relevant situation), Worth and Book found that people low in Honesty-Humility were more likely to engage in actions that ruined their opponents' experience such as attacking weaker players who did not stand a chance against them. Similarly, persons high in Conscientiousness tended to seek out to craft rare items that required a lot of work to complete (e.g., completing hour long quests to collect the required ingredients). And finally, with respect to outcome activation (i.e., the trait- and situation-contingent gains and losses), a person low in Honesty-Humility who engaged in player-versus-player activities could receive a high leaderboard score, whereas a person high in Conscientiousness could earn a lot of in-game gold by selling rare items.

Environmental and avatar choice (situation activation) and appearance (trait activation) play an important role in our design of virtual cues. For instance, Wohn and Wash (2013) found that personality traits could be inferred from the customization of a virtual city

in an internet game, reflecting findings by Gosling et al. (2002) on inferring personality from offices and living spaces. Furthermore, designing an avatar obviously overlaps with physical appearance customization, and thus, studies on avatar design yielded valid cues of personality (Bélise & Bodur, 2010; Fong & Mar, 2015). Together, these studies suggest that it is possible to develop a valid gamified personality assessment instrument on the basis of virtual cues.

The gamified instrument we developed consists of a number of such virtual cues.[2] In the instrument several pictures are presented that take virtual cues as input (e.g., various outfits that differ in how plain versus glamorous they look; see Figure 3.1) and the choice (of such an outfit, for example) of participants as output. The virtual cues instrument, here developed, may overcome some of the above-mentioned limitations of real-world personality cues. First, in a gamified setting a variety of topics—such as virtual offices, avatars, and virtual vehicles—can be included, securing a broad range of trait cues. Second, the number of cues can be controlled and scored automatically. Third, a gamified instrument does not require a full-fledged game and can be used as a stand-alone tool.

A Practical Advantage of Virtual Cues

A practical advantage of the gamified virtual cues instrument is that it might be somewhat more difficult to fake than a standard self-report personality instrument. In high-stakes selection assessments job candidates are able to selectively inflate their scores on relevant personality traits on personality inventories (Birkeland, Manson, Kisamore, Branninck, & Smith, 2006) and such faking effects are pronounced for socially desirable traits such as Honesty-Humility (Anglim, Morse, De Vries, MacCann, & Marty, 2017). Because responses to virtual cues may rely more strongly on implicit processes, they are expected to be less open to social desirability than self-report inventories and therefore more difficult to fake.

Figure 3.1: Example virtual cues item from Study 3.1 with the male avatars on the top row and female avatars on the bottom row. Outfit on the left indicative of high Honest-Humility and the outfit on the right being indicative of low Honesty-Humility.

Present Research

The first objective of our research was to demonstrate that Honesty-Humility can be inferred from virtual cues by showing an association between virtual cues and self-reported Honesty-Humility (Study 3.1). We used these cues to develop and validate a gamified personality assessment instrument by showing that a scale consisting of virtual cues is related to self-reported Honesty-Humility (Study 3.2). Subsequently, we tested whether the gamified instrument was more difficult to fake than a traditional personality inventory (Study 3.3). These studies provide a theoretical contribution to the literature by identifying personality cues potentially related to Honesty-Humility in the physical world. Furthermore, these studies

make a practical contribution by demonstrating the usefulness of virtual cues in personnel selection.

Study 3.1

The goal of Study 3.1 was to demonstrate that Honesty-Humility can be inferred from virtual cues such as avatars, virtual offices, and vehicles. These virtual cues were validated by correlating them with answers on a self-reported personality inventory. Study 3.1 was part of a larger exploratory study examining alternative Honesty-Humility measures, however, for ease of comprehension, we do not report these additional tasks (e.g., an effort task).[3]

We also explored whether more fine-grained approaches for the virtual cues would be useful by letting participants choose hypothetical car modifications (see materials for examples). We separated car modifications in three groups: interior-, exterior-, and safety modifications. Beforehand, we hypothesized that safety modifications are correlated to Emotionality because such modifications likely appeal to fearful people. Exterior and interior modifications were expected to be related to Honesty-Humility because they signal status and luxury.

Methods

Participants and procedure

In this study, 104 students (M_{age} = 21.30; SD_{age} = 3.93; 34 men) completed an internet questionnaire on a lab computer and performed three behavioral tasks (e.g., a branching situational judgment test) for course credit or monetary compensation.

Materials

HEXACO-100. Participants completed the Dutch HEXACO-100 (De Vries, Ashton, & Lee, 2009), which consists of 16 items for each of the six broad personality factors and four items for an interstitial facet scale named Altruism. Responses were provided on a five-point Likert scale (1 = *strongly disagree* and 5 = *strongly agree*). Quality checks were performed

following Lee and Ashton (2018); no responses were considered to be of low quality, thus no data was discarded.

***Virtual cues*.** Below all virtual cues are described. Note that for all virtual cues the answers and expenditures were coded so that higher scores reflected higher Honesty-Humility.

Avatars. Using the Sims 4 (Electronic Arts, 2014b) six male and six female avatars were created with outfits differing in their luxurious feeling (see Figure 3.1). Participants selected one out of these six avatars matched to their gender.

Offices. Similar to the avatars, four offices were created using Sims 4 (Electronic Arts, 2014b) that differed in their lavishness. Again, participants selected one of these four offices.

Cars. Seven different cars that differed in their luxurious appeal were created on the website http://www.3dtuning.com/. Participants selected one of these cars.

Car customization. After choosing a car, participants could spend €2,000.- on customizing the car. These choices were imaginary and not visualized. All modifications together costed €3,000.-. Participants could choose not to spend any or all of their budget. All modifications were a-priori divided into three categories: interior- (e.g., front seat heating), exterior- (e.g., vanity plate), and safety modifications (e.g., improved brakes) and the maximum budget for each category was €1,000.-.

Watches. Ten pictures of watches were created, each accompanied with a price tag, ranging from €50.- to €250.-. Participants indicated which watch they liked the most and which they would choose for their work anniversary. Participants could also refrain from choosing a watch for this second question. Three outcomes were calculated, the relative value of the watch they liked the most, the relative value of the watch they chose, and the relative difference between the value of the liked and the chosen watch.

Results and Discussion

Study 3.1 investigated whether it was possible to infer Honesty-Humility from the virtual cues by correlating each virtual cue with the HEXACO factor scales (See Table 3.1). Additionally, separate relative weights analyses (Johnson, 2000) were performed for each of the virtual cues using the six HEXACO factor scales plus gender and age because of their potential roles in the explanation of personality traits and virtual cues (See Table 3.2). Overall, the results showed Honesty-Humility was related to avatar choice (male avatars: r = .36, p = .034; female avatars: r = .34, p = .004)[4] and office choice (r = .32, p = .001). That is, people scoring lower on Honesty-Humility preferred an avatar with more expensive looking clothes and a larger office space, whereas people who scored higher on this trait preferred cheaper looking clothes and a smaller office space. Additionally, Honesty-Humility primarily contributed to the explanation of avatar and office choices. However, not all virtual cues were related to Honesty-Humility: car choice and all watch choices were more strongly related to other traits. Nonetheless, these six virtual cues were theoretically developed to assess Honesty-Humility so we considered the standardized scores of these cues as a scale. These six virtual cues had low reliability (α = .45). This scale was significantly related to Honesty-Humility (r = .33, p = .001), Emotionality (r = .26, p = .007), and gender (r = -.30, p = .002; with F=0, M=1).

Considering the more fine-grained car modifications, the interior modifications were significantly correlated with Honesty-Humility (r = .38, p < .001), Emotionality (r = .24, p = .014), and Conscientiousness (r = .22, p = .022), whereas the exterior modifications were significantly correlated with Honesty-Humility (r = .24, p = .014) and Extraversion (r = -.24, p = .014). The relative weights analyses revealed that interior modifications were mainly related to Honesty-Humility, whereas exterior modifications were related mostly to Extraversion, and safety modifications were unrelated to any of the personality traits.

Table 3.1

Correlations between the demographic variables, HEXACO personality traits, and the virtual cues.

Variable		*M*	*SD*	1	2	3	4	5	6	7	8
1.	Gender	.33	.47	-							
2.	Age	21.40	3.93	.09	-						
3.	H	3.32	.70	-.37**	.03	-					
4.	E	3.14	.66	-.51**	-.11	.26**	-				
5.	X	3.55	.60	.13	-.16	-.17	-.04	-			
6.	A	3.05	.57	-.13	.04	.47**	.16	.01	-		
7.	C	3.35	.55	-.21*	.01	.28**	.22*	-.19*	.01	-	
8.	O	3.45	.61	.15	.05	-.09	-.11	.09	-.13	-.17	-
9.	Avatar	3.13	1.63	-.08	.06	.35**	.19*	-.06	.17	-.02	-.10
10.	Office	1.88	.97	-.21*	-.11	.32**	.24*	-.01	.18	.03	-.05
11.	Car	3.39	1.19	-.29**	-.02	.01	.26**	.04	.02	-.00	-.29**
12.	CA: Internal	425.00	290.88	-.30*	.09	.38**	.24*	-.14	.01	.22*	-.05
13.	CA: External	817.31	223.15	-.20*	.04	.24*	.16	-.24*	.15	.08	-.07
14.	CA: Safety	468.27	371.37	-.05	.06	.04	-.02	-.01	-.01	-.06	-.13
15.	Watch: liking	2.84	1.50	-.07	.02	.17	.02	-.11	.08	-.01	.17
16.	Watch: choice	2.65	1.23	-.17	-.08	.15	.05	-.11	.08	.11	.27**
17.	Watch: difference	.18	1.43	-.12	-.10	.01	.04	-.02	.01	-.11	.13

Table 3.1 continued

Variable		9	10	11	12	13	14	15	16
10.	Office	.26**	-						
11.	Car	-.04	.17	-					
12.	CA: Internal	.19	.22*	.16	-				
13.	CA: External	.26**	.37**	.09	.17	-			
14.	CA: Safety	.15	.17	.05	.30**	.08	-		
15.	Watch: liking	.19	.10	.10	.03	.08	.02	-	
16.	Watch: choice	.04	.18	.08	-.01	.10	-.15	.46**	-
17.	Watch: difference	-.12	.10	.00	-.04	.04	-.18	-.37**	.65**

Note. For gender F=0, M=1; H = Honesty-Humility; E = Emotionality; X = Extraversion, A = Agreeableness; C = Conscientiousness; O = Openness to Experience; CA = car adjustments; The virtual cues variables were coded that higher scores reflect more Honesty-Humility; * = $p < .05$; ** $p < .01$.

Table 3.2

Relative weights analyses of the virtual cues, the demographic variables and the HEXACO personality traits.

Variable	Gender	Age	H	E	X	A	C	O	R^2
Avatar	.006	.004	**.106**	.030	.002	.012	.011	.008	.18
	3.4%	2.2%	**59.2%**	16.7%	1.0%	6.5%	6.2%	4.7%	
Office	.016	.009	**.071**	.033	.001	.013	.003	.001	.15
	10.6%	6.4%	**48.3%**	22.6%	.5%	9.1%	1.8%	.6%	
Car	.056	.000	.004	.045	.004	.001	.004	**.078**	.19
	29.1%	.2%	2.2%	23.4%	2.0%	.5%	2.1%	**40.7%**	
CA: Interior	.043	.010	**.110**	.027	.007	.015	.021	.001	.23
	18.4%	4.3%	**47.1%**	11.4%	2.9%	6.6%	9.0%	.3%	
CA: Exterior	.018	.001	.028	.012	**.047**	.012	.001	.002	.12
	14.8%	.9%	23.1%	10.1%	**38.8%**	9.8%	1.0%	1.4%	
CA: Safety	.002	.004	.002	.001	.000	.002	.007	**.018**	.04
	6.5%	11.6%	6.7%	2.3%	.4%	4.4%	18.3%	**49.8%**	
Watch: liking	.004	.000	.023	.001	.011	.004	.002	**.034**	.08
	4.6%	.2%	29.6%	.7%	13.9%	5.7%	1.9%	**43.3%**	
Watch: choice	.027	.009	.017	.002	.016	.004	.019	**.079**	.17
	15.5%	5.0%	10.0%	1.4%	8.9%	2.5%	11.0%	**45.6%**	
Watch: difference	**.016**	.010	.001	.002	.002	.000	.014	**.019**	.06
	24.6%	16.2%	1.3%	2.6%	2.9%	.4%	22.2%	**24.6%**	

Note. H = Honesty-Humility; E = Emotionality; X = Extraversion, A = Agreeableness; C = Conscientiousness; O = Openness to Experience; On the first row for each virtual cue the raw relative weight is reported, on the second row the relative weight of each variable is reported in percentages of the total explained variance. The highest parameter is highlighted in bold.

Study 3.2

Study 3.1 demonstrated that Honesty-Humility can be assessed with virtual cues. Furthermore, the results of the car modification task suggest that more fine-grained components of virtual cues reflect different traits. For example, when designing an avatar, choices in jewelry, shoes, and clothes may reflect different personality traits.

Based on the results of Study 3.1, we developed a set of more fine-grained items that were expected to form, when combined, a valid virtual cues scale of Honesty-Humility. Again, items were validated by relating them to the self-reported Honesty-Humility.

Methods

Participants and procedure

Using Amazon Mechanical Turk (MTurk), 219 American participants were recruited for monetary compensation. Participants completed the HEXACO-100 (Lee & Ashton, 2018) and the virtual cues instrument (in a randomized order). Participants were prevented from filling out the questionnaire on a phone or tablet to guarantee clear visibility of the pictures. Based on the exclusion criteria (see the materials section below) data of 23 participants were discarded, which resulted in a sample of 196 participants (M_{age} = 38.56; SD_{age} = 11.41; 110 men).

Materials

HEXACO-100. Participants completed the HEXACO-100 without the Altruism facet scale (i.e., the first 96 items of the HEXACO-100, see Lee & Ashton, 2018). Instead, four Infrequency items (e.g., *"I have never used a telephone"*; Fekken, Holden, Jackson, & Guthrie, 1987) were included at specific places in the HEXACO-100 to detect noncompliant responses. Data of respondents who did not answered with "strongly disagree" or "disagree" on two or fewer of these questions were not included in the final dataset (see Barends & De

Vries, 2019 for more details). Again, Lee and Ashton's (2018) quality checks were performed (see also Barends & De Vries, 2019 for more information).

Virtual cues. A total of 71 virtual cues were developed based on the computer games Sims 3 (Electronic Arts, 2009), Sims 4 (Electronic Arts, 2014b), SimCity (Electronic Arts, 2014a), and Grand Theft Auto V (Rockstar Games, 2013). Furthermore, items were developed based on images that were not copyrighted and labeled for re-use and faces from the FaceGen database (Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013; Todorov & Oosterhof, 2011). In Study 3.2 all items consisted of four pictures. For half of the items the pictures were presented from high to low Honesty-Humility and for the other half the order was reversed. We developed 36 items for Honesty-Humility. The other 35 items—which are not reported here—were developed for the other traits (see footnote 3).

Results and Discussion

The goal of Study 3.2 was to develop a valid and reliable virtual cues Honesty-Humility scale. We found that the 36 virtual cues formed a scale with acceptable reliability (α = .78) and this scale showed convergent validity with self-reported Honesty-Humility (r = .41, p < .001). Based on relative weights analyses for each of the 36 items with the six self-reported HEXACO scores, we selected 18 items with the highest importance weights for Honesty-Humility. These 18 items together formed a reliable scale (α = .72) and this scale was also significantly correlated with self-reported Honesty-Humility (r = .49, p < .001; see Table 3.3).

To investigate the extent to which the final virtual cues scale reflects Honesty or Humility the four facets were included in a relative weights analysis (Johnson, 2000). The results revealed that the model accounted for R^2 = .29 and most variance was due to the Humility facets (Greed Avoidance, 46.6%; Modesty, 32.4%), whereas the Honesty facets explained less variance in the virtual cues scale (Fairness, 16.4%; Sincerity, 4.5%).

Table 3.3

Correlations between the demographic variables, HEXACO personality traits, and the virtual cues Honesty-Humility scale in Study 3.2.

Variable		*M*	*SD*	1	2	3	4	5	6	7	8	9
1.	Gender	.56	.50	-								
2.	Age	38.56	11.41	-.24*	-							
3.	H	3.57	.71	-.22*	.29**	-						
4.	E	3.09	.68	-.40*	.07	.00	-					
5.	X	3.21	.84	.15*	.09	.07	-.30**	-				
6.	A	3.32	.66	-.09	.09	.37**	-.17*	.41**	-			
7.	C	3.85	.52	-.06	-.08	.25**	-.17*	.38**	.24**	-		
8.	O	3.56	.71	.07	.09	.20**	-.06	.32**	.15*	.27**	-	
9.	Full VC-H	2.45	.31	-.11	.24**	.40**	.13	-.25**	.08	-.14*	.03	
10.	Final VC-H	2.48	.39	-.13	.27**	.49**	.13	-.16*	.15*	-.11	.00	.91**

Note. For gender F=0, M=1; H = Honesty-Humility; E = Emotionality; X = Extraversion, A = Agreeableness; C = Conscientiousness; O = Openness to Experience; VC = Virtual cue; * = $p < .05$; ** $p < .01$.

Study 3.3

Study 3.2 described the development of a gamified scale to measure virtual cues of Honesty-Humility. In Study 3.3, because the content of the gamified scale is based on unobtrusive cues, we expected the instrument to be more difficult to fake than the HEXACO-100. We assessed the fakeability in a simulated high-stakes selection situation in which the ideal candidate would score high in Honesty-Humility. Prior work has found that in such simulations participants selectively inflate their scores on items (and thus also their trait scores) when they perceive the items as relevant for the job (e.g., Shoss & Strube, 2011).

Methods

Participants and procedure

A sample of 220 American participants were recruited via MTurk and were offered monetary compensation. Participants from Study 3.2 were excluded from partaking in Study 3.3. Based on the same exclusion criteria as in Study 3.2, the data of 29 participants were

discarded. This resulted in a sample of 191 participants (M_{age} = 40.83; SD_{age} = 11.85; 68 men). Participants were randomly assigned to one of four conditions in which they either completed both, only one, or none of the scales under faking instructions (and the other instrument(s) under honest instructions). Within each condition, instruments were completed in a counterbalanced order. After completion of each instrument, a test perception inventory was completed.

Materials

HEXACO-100. The same 96 items of the HEXACO-100 were used as in Study 3.2.

Virtual cues. The 18-item virtual cues instrument as reported in the results of Study 3.2 was used together with 26 filler items that covered the other five traits (results of these five scale are not reported here—see footnote 3). Reliabilities of the Honesty-Humility virtual cues were α = .65 (in the honest condition) and α = .62 (in the faking condition).

Job description. A job description was created that required candidates to be high on Honesty-Humility and that the best "candidate" received a bonus of $25.-. However, a warning signal was included that a psychologist selected the best "candidate" and might be wary of responses that were too good to be true. The bonus was set at a high level compared to the participation fee ($2.-) to motivate participants to perform well (as in a real high-stakes selection). Note that we counterbalanced the order of instrument administration also in the two conditions in which only one instrument was completed under the faking instructions. Therefore, half of the participants in these conditions were administered the faking condition before the honest condition.

Manipulation check. As a manipulation check, we included a self-developed six-item faking attempt scale as part of the test perception questionnaire. An example item of the faking attempt scale is *"I have filled in the questionnaire honestly"*. Answers were given on a

five-point Likert scale (1 = *strongly disagree* and 5 = *strongly agree*). Across the four conditions, *α's* for this attempts at faking scale ranged between .55 and .82.

Results and Discussion

In Study 3.3, we attempted to replicate our findings from Study 3.2 by correlating the virtual cues scale with self-reported Honesty-Humility. The virtual cues scale under honest instructions was significantly correlated with Honesty-Humility under honest instructions (r = .30, p = .037). Under faking instructions the virtual cues scale also showed convergent validity with both honest (r = .34, p = .018) and faked self-reported Honesty-Humility (r = .38, p = .009) (See Table 3.4).

To determine whether the job description had the intended effect we compared the faking attempt scale between conditions in which the self-reported HEXACO was administered with faking instructions (M = 2.21; SD = .84) and honest instructions (M = 1.78; SD = .52) using an independent samples t-test. The difference between the two conditions was significant, $t(155.91) = 4.44$, $p < .001$, $d_s = .62$ (equality of variances cannot be assumed). Similarly, the faking attempts were significantly higher for the virtual cues of those who received faking instructions (M = 2.09; SD = .78) than those who received honest instructions (M = 1.88; SD = .54), $t(163.35) = 2.13$, p = .034, d_s = .31 (equality of variances cannot be assumed). These results indicate that manipulation was successful for both instruments.

Faking. The fakeability of both the HEXACO-100 and the virtual cues instrument were investigated with one-way ANCOVA's. The honest and faking conditions of the instruments and the order of administration were included as between factors and age and gender were included as covariates. As expected, the mean score of self-reported Honesty-Humility in the faking condition (M = 3.81, SD = .66) was significantly higher than in the honest condition (M = 3.62, SD = .67), $F(1, 185) = 5.61$, p = .019, η_p^2 = .029 (See Table 5). However, contrary to our expectations, the mean score of the virtual cues Honesty-Humility

scale in the faking condition (M = 2.67, SD = .32) was also significantly higher than in the honest condition (M = 2.50, SD = .35), $F(1, 185) = 12.84$, $p < .001$, $\eta_p^2 = .065$ (See Table 5). Furthermore, there was no significant interaction for both instruments between the condition and the order of administration (See Table 3.5), so there was no spill-over effect of receiving the faking instructions before the honest instructions.

These effects suggest that the virtual cues Honesty-Humility scale may be actually somewhat *easier* to fake than self-reported Honesty-Humility. In order to investigate this possibility the virtual cues scale was recoded to a five-point scale. In a mixed MANCOVA, the virtual cues and HEXACO-100 Honesty-Humility scales were included as within factors, the instructions (faking vs. honest) for each instrument and the order of administration as between factors, and gender and age as covariates. The multivariate analyses of the interaction between the instruments and the instructions for both instruments showed that the strengths of the effects did not differ, $F(1, 181) = 2.36$, $p = .126$, $\eta_p^2 = .013$.[5] That is, the Honesty-Humility virtual cues scale was not significantly easier to fake than the self-report scale.

General Discussion

The present research demonstrates that virtual cues can be used to infer Honesty-Humility. First of all, this research shows that it is possible to construct a reliable virtual cues scale to measure Honesty-Humility. Second, simulated high-stakes selection situations did not have a strong influence on the convergent validity of the Honesty-Humility virtual cues scale. The results of this study do not support our contention that virtual cues are more difficult to fake. Both the HEXACO-100 and the virtual cues instrument could be faked and to similar extent. Of course, the warning signal included in the study may have limited the extent of faking as has been found in earlier work (McFarland, 2003; O'Neill et al., 2013). Thus, the upper bound of faking the virtual cues Honesty-Humility scale is not yet known.

Table 3.4

Correlations between the Honesty-Humility virtual cues scales and the HEXACO-100 in the honest and faking conditions.

Variable	*M*	*SD*	VC-H (honest)	VC-H (faked)
Gender	0.64	.48	-.07	.17
Age	40.83	11.85	.21*	.06
H (honest)	3.62	.67	.30*	.34*
E (honest)	3.16	.64	-.02	.29*
X (honest)	3.43	.63	-.09	-.01
A (honest)	3.40	.54	.15	.19
C (honest)	3.99	.53	-.07	-.01
O (honest)	3.67	.63	-.26	-.04
H (faked)	3.81	.66	.32*	.38**
E (faked)	3.08	.61	-.21	-.11
X (faked)	3.59	.72	-.13	-.20
A (faked)	3.46	.67	.16	.25
C (faked)	4.19	.51	-.01	-.04
O (faked)	3.70	.66	-.03	.04

Note. H = Honesty-Humility; E = Emotionality; X = Extraversion, A = Agreeableness; C = Conscientiousness; O = Openness to Experience; VC = Virtual cues; * = $p < .05$, ** $p < .01$; $n = 49$ for the correlations between the HEXACO-100 (honest) and the VC-H (honest); $n = 47$ for the correlations between the HEXACO-100 (honest) and the VC-H (faked); $n = 49$ for the correlations between the HEXACO-100 (faked) and the VC=H (honest); $n = 46$ for the correlations between the HEXACO-100 (faked) and the V-HC (faked).

Table 3.5

Results of the two ANCOVA's of the HEXACO-100 and the virtual cues instrument when comparing the honest and instructed faking conditions.

	HEXACO			Virtual Cues		
	F	*P*	η_p^2	*F*	*p*	η_p^2
Gender	8.28	.004	.04	.25	.618	.00
Age	16.67	.000	.08	3.84	.052	.02
Order	.11	.747	.00	.36	.552	.00
Faking instructions	5.61	.019	.03	12.84	.000	.07
Order X instructions	.01	.931	.00	.83	.362	.00

Note. DF of all models (1, 185)

Theoretical Implications

We demonstrated that Honesty-Humility can be inferred from virtual cues in a wide variety of domains such as avatars, virtual offices, and vehicles because the choices people made in these domains were significantly correlated to self-reported Honesty-Humility. More specifically, people low on Honesty-Humility are more likely to choose more luxurious and expensive-looking versions of these cues whereas people high on this trait choose more modest looking cues. Not surprisingly, these cues are more strongly related to the humility facets than the honesty facets. These findings suggest that Honesty-Humility can also be inferred from real-world personality cues. For instance, the investigation of Vazire et al. (2008) on personality cues of narcissism may have already identified several cues related to Honesty-Humility such as wearing a lot of makeup or wearing expensive clothes.

We should warn, however, that in line with the "mapping principle" (Williams, 2010), our findings in a gamified setting cannot be taken as direct evidence of real-life Honesty-Humility cues. This principle suggests that behaviors in the real- and virtual world may not necessarily be equivalent, even though they may share a strong surface resemblance. For instance, the number of people killed in a video game may be an indicator of Conscientiousness, whereas in the real world this is most likely related to (extremely) low Honesty-Humility. Thus, although the items were developed on a theoretical basis and although we expect many cues to accurately map on Honesty-Humility, some of the identified personality cues may map differently on traits in the real world.

Another difference between real- and virtual cues is that—in our studies—choices were restricted by only presenting several alternatives to choose from, whereas in real-life there are many more options available. Moreover, our participants could choose alternatives to which they probably had no access in real-life such as a Ferrari and a yacht. However, this latter feature can be regarded as a strength of the virtual cues approach because more cues can

be utilized than would be possible in the real world. These differences require that future studies examine which personality cues are indicative of Honesty-Humility in the real world.

An outstanding question is why virtual cues can be as easily faked as standard personality inventories. A potential explanation is that our virtual cues are more akin to identity claims than behavioral residues (Gosling et al., 2002). Arguably, behavioral residues are more unobtrusive—and indirect—personality cues than identity claims as they result from behavior (e.g., a messy office is likely the result of low Conscientiousness). On the other hand, identity claims have an intentional component (dyeing one's hair blue to be perceived as rebellious) and are less unobtrusive, and thus more fakeable. This closely aligns with research demonstrating that even in daily life self-presentation motives have an effect on how people dress, for instance to convey a specific image (Nezlek, Mochort, & Cypryańka, 2019).

Limitations and Directions for Future Research

A limitation of the current studies for practical usage of the virtual cues is that we did not investigate its predictive validity. We only obtained evidence for construct validity, thus it is unknown whether the virtual cues instrument predicts work outcomes such as counterproductive work behavior (Zettler & Hilbig, 2010). However, it should be noted that most gamified assessment tools typically fail to provide evidence of construct validity as noted by Chamorro-Premuzic, Winsborough, Sherman, and Hogan (2016).

Future research might assess other potential advantages of the virtual cues instrument. Typically, increasing the motivational potential (e.g., liking) of a system is the goal of gamification (Deterding et al., 2011). Particularly, in personnel recruitment and assessment procedures gamified tools might improve candidate reactions. Thus, future research might also like to investigate whether virtual cues instruments result in favorable applicant reactions.

Conclusion

The current set of studies indicates that virtual cues of Honesty-Humility are available in domains such as avatar design, virtual vehicle choice, and office customization. These virtual cues may be especially useful as valid and reliable indicators of Honesty-Humility in a gamified assessment tool. This novel way to measure personality can be easily incorporated in full-fledged selection assessment games as a form of 'stealthy' assessment (Shute, Masduki, & Donmez, 2010).

Notes

1. Graham et al. (2011) also list a third type of personality cue: thought and feeling regulators that are deliberate changes to an environment to regulate emotions and/or thoughts. For instance, keeping vacation pictures in ones' office to relive treasured memories. However, this category may be largely subsumed by the other two types.

2. Note that virtual cues as developed in this paper are not derived from personality traces left in a virtual environment (i.e., virtual traces) but from choices in a formal virtual environment (i.e., virtual choices). The main distinction between the two is that traces are often derived based on 'open choices' which are often more than a limited number of options as we use here. However, these virtual traces are also distinct from real-world personality traces as the system should allow for the existence of such traces. For instance, The Sims game explicitly programmed Sims' behavior to create a mess in their in-game house. Thus, the messiness of the house would be the virtual trace left behind. If these behaviors had not been included in the game, this virtual trace would not have existed.

3. Detailed methods and results of variables mentioned, but not reported, in Studies 1, 2, and 3 are available upon request.

4. Using Meng's z-scores (Meng, Rosenthal, & Rubin, 1992), we found that the correlation coefficients of men and women were not statistically different ($z = .10$, $p = .92$) and therefore the findings on the avatars were aggregated across genders.

5. Note that we did not analyze the main effects with these recoded scores as this would have confounded the type of instrument with number of response options. Therefore, we do not report the full MANCOVA.